

Optimizing the Design of an Artificial Pancreas to Improve Diabetes Management

Ashok Khanna
Georgia Institute of Technology
Cognizant Technology Solutions
ashok.khanna@cognizant.com

Olivier Francon
Cognizant AI Labs
olivier.francon@cognizant.com

Risto Miikkulainen
The University of Texas at Austin
Cognizant AI Labs
risto@cs.utexas.edu

ABSTRACT

Diabetes, a chronic condition that impairs how the body turns food into energy, i.e. blood glucose, affects 38 million people in the US alone [5]. The standard treatment is to supplement carbohydrate intake with an artificial pancreas, i.e. a continuous insulin pump (basal shots), as well as occasional insulin injections (bolus shots). The goal of the treatment is to keep blood glucose at the center of an acceptable range, as measured through a continuous glucose meter. A secondary goal is to minimize injections, which are unpleasant and difficult for some patients to implement. In this study, neuroevolution was used to discover an optimal strategy for the treatment. Based on a dataset of 30 days of treatment and measurements of a single patient, a random forest was first trained to predict future glucose levels. A neural network was then evolved to prescribe carbohydrates, basal pumping levels, and bolus injections. Evolution discovered a Pareto front that reduced deviation from the target and number of injections compared to the original data, thus improving patients' quality of life. To make the system easier to adopt, a language interface was developed with a large language model. Thus, these technologies not only improve patient care but also adoption in a broader population.

KEYWORDS

Surrogate modeling, Decision making, Diabetes, Neural Networks, Evolutionary Computation

1 INTRODUCTION

About 422 million people worldwide have diabetes, most of them in the low- and middle-income countries, and 1.5 million deaths are directly attributed to diabetes each year [20]. Diabetes leads to serious issues in the long term, such as heart conditions, vision loss, or kidney failure; it is the #1 cause of kidney failure, lower-limb amputations, and adult blindness [5]. In the USA alone, \$1 in \$4 is spent on care for diabetes, with an economic burden of e.g. \$327 Billion in 2017. Also, 277,000 premature deaths in the US were attributed to diabetes that year [7].

Of the two types, this paper focuses on Type 1 diabetes that often occurs early in life [4]. The patient's pancreas does not produce enough insulin and the patient is dependent on an external dosage of insulin multiple times a day. Recent advances in closed-loop automated insulin delivery systems offer a paradigm shift in Type 1 diabetes management [15–17]. The idea is to deliver insulin to patients through a continuous low-rate wearable insulin pump (basal insulin shots), supplemented with injections as needed (bolus insulin shots). **Blood Glucose level is measured continuously using a wearable Continuous Glucose Meter (CGM; Figure 1). These devices are available from Abbott [1], Medtronic [13], Dexcom [6] and other**

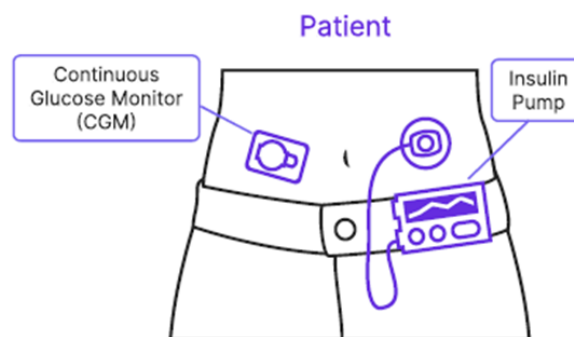


Figure 1: An artificial pancreas, i.e. a closed loop system to manage diabetes. The patient has attached a Continuous Glucose Monitor and a pump for Insulin delivery. The goal is to adjust the pump rate to keep the glucose level constant. If the pump is not sufficient, insulin injections can be added as well, but the goal is to try to avoid them. [Figure from 16]

manufacturers [10], and can improve the patients' quality of life significantly [2].

The goal is to keep the A1c, i.e. the three-month average percentage of glucose-bound red blood cells below 6.5% [8] and randomly timed glucose level measurements in the range of 70 mg/dl to 180 mg/dl [16]. **A variety of factors contribute to blood glucose level [9], which makes this task difficult. As a matter of fact, a patient may make about 100 decisions and spend approximately 58 mins per day managing their diabetes [3], which amounts to approximately 4% of their life.**

The goal of this paper is to optimize this decision-making process. The idea is to use this historical data to learn what worked for patients in the past, and use these insights to improve patient care. More specifically, the system learns to recommend carb intake, basal rate, and bolus dose so that the blood glucose level is as stable as possible, and the bolus injection dosage are minimized.

Data collection with all the potential factors affecting diabetes is challenging. **Moreover, at least 14 days of glucose levels data is needed to get a high correlation with three-month averages; short term data is not sufficient [18]. Therefore, our approach is based on a dataset of 729 hourly samples, or about 30 days of data, from a single representative patient in the DiaTrend dataset [16].** The approach involves training two models: The predictive model uses five hours of past data to predict the glucose level deviations in the next hour. The prescriptive model uses the same five-hour input to

recommend carb, basal, and bolus dosage such that the deviations and the necessary bolus dose are minimized.

These results are encouraging. When limited to the same range of actions as the original dataset, the system results in more stable blood glucose levels and fewer injections, demonstrating that better care is possible even within the current practice. When the range of actions was extended to larger values, even better results were obtained, suggesting that new and better approaches are possible as well.

The system also allows modifying the recommendations at will and seeing how well they would work. This approach gives much more flexibility to the patient to make choices about their daily life—be it food intake, basal rate, or bolus injections. For instance, if a patient plans to attend a party and expects the carb intake to be higher than his or her normal daily routine, the patient can check what the expected blood glucose level deviation would be and make a decision about bolus injections accordingly. To further improve the adoption of this tool, a user interface with LLM was developed where user can enter questions in prompt to interact with the model. In this manner, AI technologies can not only improve diabetes management, but also make the improvements more broadly accessible.

2 BACKGROUND

A patient’s response to insulin and carb intake varies depending on current blood glucose level and their diabetes profile. Given the wide range of these parameters, it is not possible to consider all the combinations and prescribe precise treatment ahead of time. **A practical solution, then, is to check the blood glucose level frequently and change the treatment to keep the blood glucose level in an acceptable range.** This approach is called artificial pancreas. The technology adopted in this paper to optimize it is called Evolutionary Surrogate-assisted Prescription.

2.1 Artificial Pancreas

Artificial Pancreas, i.e. a closed loop insulin delivery mechanism (Figure 1) consists of a Continuous Glucose Monitor (CGM) that is used to check blood glucose level periodically (e.g. in 5 min intervals), and an insulin pump that patients wear to deliver continuous insulin (i.e. basal shots) at a low flow rate. In addition, the patients may occasionally need insulin injections as well (i.e. a bolus shots). It is possible to learn how the patient responds to insulin and carb intake, and adjust these rates accordingly by hand.

However, there are opportunities to optimize these decisions further. It may be possible to explore a wider range of carb intake, basal rates, and bolus dosages to make them more effective; offer alternative treatments depending on patient preferences; decide whether a bolus injection will be needed overnight (when patient is sleeping) and plan accordingly; etc. Such flexibility can empower diabetic patient to take control of their diabetes and enjoy life with fewer restrictions from their condition. While currently these strategies are found manually through trial and error, it may be possible to learn good strategies from historical data. A possible technology for doing that will be outlined in the next section.

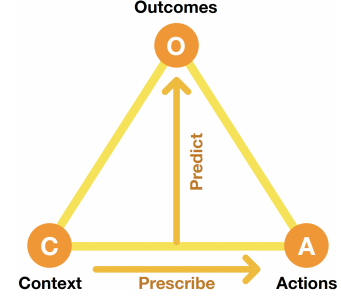


Figure 2: The ESP Decision Optimization Method. A predictor is trained with historical data on how given actions in given contexts led to specific outcomes. It is then used as a surrogate in order to evolve prescriptors, i.e. neural networks that implement decision policies resulting in the best possible outcomes.

2.2 Evolutionary Surrogate-assisted Prescription

Evolutionary Surrogate-assisted Prescription [ESP; 11] is an approach for optimizing decision-making in a variety of domains (Figure 2). The main idea is that a decision policy can be represented as a neural network, or a set of rules, and a good policy can be discovered through population-based search, i.e. using evolutionary computation techniques. However, each candidate must be evaluated, which is difficult to do in many real-world applications. Therefore, a surrogate model of the world is learned from historical data, predicting how good the resulting outcomes are for each decision in each context.

More formally, given a set of possible contexts \mathbb{C} and possible actions \mathbb{A} , a decision policy D returns a set of actions A to be performed in each context C :

$$D(C) = A, \quad (1)$$

where $C \in \mathbb{C}$ and $A \in \mathbb{A}$. For each such (C, A) pair there is a set of outcomes $O(C, A)$, and the Predictor P_d is defined as

$$P_d(C, A) = O, \quad (2)$$

and the Prescriptor P_s implements the decision policy as

$$P_s(C) = A, \quad (3)$$

such that $\sum_{i,j} O_j(C_i, A_i)$ over all possible contexts i and outcome dimensions j is maximized (assuming they improve with increase). It thus approximates the optimal decision policy for the problem. The predictor can be learned from historical data on CAO triples. In contrast, the optimal actions A for each context C are not known, and must therefore be found through search.

ESP was first evaluated in reinforcement learning tasks such as function approximation, cart-pole control, and the flappybird game, and found to discover significantly better solutions, find them faster, and with lower regret than standard approaches such as direct evolution, DQN, and PPO. Its power comes from automatic regularization and automatic curricular learning: When predictors are learned at the same time as the prescriptors, each prescriptor is evaluated against multiple predictors; as the predictors improve,

they provide more refined evaluations. ESP is thus a powerful mechanism for learning policies in time-varying domains. In this paper, it will be applied to the task of optimizing carb intake, basal rate, and bolus injection dosage for diabetes patients.

3 THE DIABETES MANAGEMENT PROBLEM

We applied ESP to prescribe carb intake, basal rate, and bolus injection dosage in order to minimize the deviation of the patient’s blood glucose level from the mid-point of the acceptable range, and to minimize the need for giving bolus injections. A predictive model was built based on historical data and a prescriptive model to optimize the outcomes. The data sources are described first, followed by the problem definition including how the data are used as the context, action, and outcome variables.

3.1 Data

The project was based on data from a single patient extracted from the DiaTrend dataset [16], available on demand. It contains three main subsets of data including Continuous Glucose Monitoring (CGM) readings, basal rate, and bolus injection dosage. Also, it captures information on patients’ carb intake as well as carb to insulin ratio over a period of more than 30 days.

The data was first aligned along a single timeline in 10 second increments. It was then aggregated into 15-minute and 60-min intervals to make decision-making process realistic. **The 60-min one was eventually used because it offers the most flexibility to the patient.** The aggregation was done for each of the variables as follows:

- Carb intake: Added all recorded carb intake values for the 60 min period;
- CGM reading: Used the last reading in the 60 min window; if it was missing, then interpolated linearly between the nearest recorded ones;
- Bolus dosage: Added all recorded bolus dosages over the 60 min period; and
- Basal rate: Multiplied the recorded rates and their durations over the 60 min period and added them up to get an hourly rate.

The resulting dataset consists of 729 hourly samples.

3.2 Decision-making Problem

Applying ESP to this domain consists of solving two challenges:

Prediction: What are the outcomes of the decision maker’s actions? That is, what is the impact on blood glucose level deviation when a decision maker takes an action for carb intake, basal rate, and bolus dosage at the current blood glucose level?

Prescription: What are the actions that result in the best outcomes? That is, is it possible to find better combinations of actions than those already in the dataset? Is it possible to do better by extending the actions with larger or smaller values than those in the dataset? How do those actions depend on the state of the patient and past decisions?

The primary goal of the artificial pancreas is to maintain a constant blood glucose level close to the midpoint of an acceptable range. However, it is also desirable to manage this goal with as few

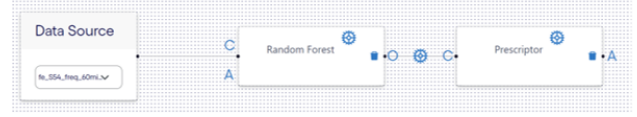


Figure 3: Building blocks for the ESP approach to managing diabetes. The image shows a screenshot of the ESP software tool that allows designing the system from pre-existing building blocks. Leftmost is the data block; it inputs data from a csv file where the samples consisting of Context, Actions and Outcomes are specified. The data is fed to a Random Forest block, trained to predict deviations from the optimal blood glucose levels. This block provides input to the Prescriptor block, which evolves a neural network to prescribe actions that result in optimal outcomes for each context.

bolus injections as possible. Injections are distracting and invasive even in the best of circumstances, but some patients may not even be able to manage insulin injections without support from caregiver. Therefore, bolus injection dosage was not only used as an action variable, but also as a cost outcome.

Accordingly, the context, action, and outcome variables for the ESP formulation were:

Context describes the problem the decision maker is facing, i.e. historical data on the patient. More specifically it consists of five previous values (i.e. spanning the previous five hours) of:

- blood glucose level (mg/dL);
- bolus injection dosage (international insulin units);
- basal rate (units/hr); and
- carb input (grams).

Actions represent the choices the decision-maker faces, i.e. what they can do at this point to control the blood glucose level of the subject. They consist of three possible recommendations:

- inject a dose of bolus insulin (a continuous value within a given range, units);
- basal rate (a continuous value within a given range, units/hr); and
- carb intake (a continuous value within a given range, grams).

Outcomes consisted of minimizing over the next 60 minutes:

- deviation of the blood glucose level from the optimal mid-point of the healthy range (mg/dL); and
- bolus injection dosage (units).

These variables define a decision-making problem that can be solved using ESP, as will be described next.

4 MODELS

The system consists of the predictor, trained with supervised learning on the historical data, and the prescriptors, trained through evolution.

Prediction: Given the context and actions that were performed, the predictive model estimates the blood glucose level deviation outcome. Any predictive model can be used in this task, including a neural network, random forest, or linear regression. As shown in the screenshot in Figure 3, a random forest was used in the current experiments. As usual, the model is fit to the existing historical

data and evaluated with left-out data. Note that the bolus injection outcome does not need to be predicted because it is determined by the prescribed bolus action directly.

Prescription: Given the context, the prescriptive model suggests actions that optimize the outcomes. The model has to do this for all possible contexts, and therefore it represents an entire strategy for optimal diabetes management. The strategy can be implemented in various ways, including sets of rules or neural networks. The approach in this paper is based on neural networks.

The optimal actions are not known, but the performance of each candidate strategy can be measured using the predictive model. Therefore the prescriptive model needs to be constructed using search techniques. Standard reinforcement learning methods such as PPO and DQN are possible; the experiments in this paper use evolutionary optimization, i.e. conventional neuroevolution [19]. As in prior applications of ESP [11, 14], the network has a fixed architecture of two fully connected layers; its weights are concatenated into a vector and evolved through crossover and mutation.

5 EXPERIMENTAL SETUP

A Random Forest model was trained with scikit-learn to predict the blood glucose level deviation. The dataset was split into an 80% training set and a 20% test set. The forest consisted of 100 decision trees with unrestricted maximum depth; at each split, a random subset of n features was considered (where n is the square root of the total number of features). Accuracy was measured in mean absolute error (MAE).

Two prescription experiments were run: In the first, the range of actions was limited to the same range as observed in the dataset: Carb intake was within [0..112] grams, basal rate within [0..0.45] units/hr, bolus injections within [0..28] units. The idea is that this range is already familiar to the patient, and we are optimizing simply to take advantage of familiar procedures better. However, since the dataset consisted of only one patient and for a limited time, not all possible variation of the variables was included in it. Therefore, in the second experiment, we widened the range for the action variables significantly: Carb intake was within [0..300] grams, basal rate within [0..1] units/hr, bolus injections within [0..200] units. Provided the predictive models can be trained to generalize well to new values, such an expansion helps explore solutions that may not have been captured within the given dataset, and to explore solutions that may not have been tried yet, leading to potential new insights.

The prescriptor network was implemented with Keras/Tensorflow. The input consisted of 20 float values and output of three float values with sigmoid activation. The hidden layer contained 16 units with tanh activation. During evolution, prescriptor candidates were evaluated on full dataset. Evolution was run for 50 generations with the following parameters:

- nb_elites: 5
- mutation_type: gaussian_noise_percentage
- nb_generations: 50
- mutation_factor: 0.1
- population_size: 100
- parent_selection: tournament
- initialization_range: 1

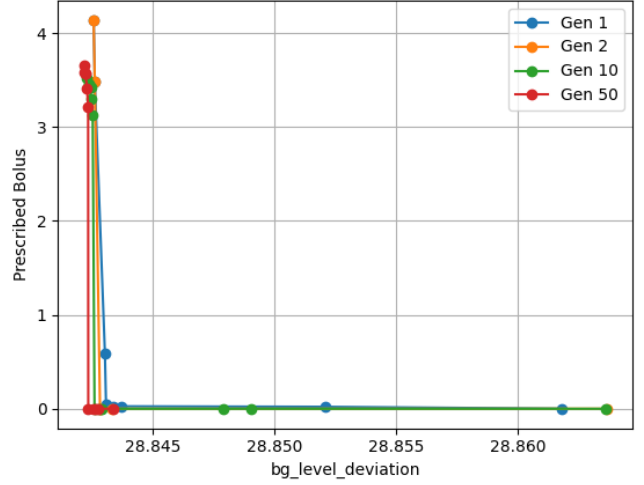


Figure 4: Progress of evolution towards better prescriptors to optimize both outcomes. Select generations (1, 2, 10 and 50) are shown for readability. Model optimizes both outcomes, i.e. minimizes value for blood level deviation as well as the bolus dosage; the plot shows a Pareto front of these tradeoffs. As evolution progresses, the front is pushed towards lower left. Interestingly, it will eventually discover solutions where the prescribed bolus dose is zero, i.e. solutions where bolus injections are eliminated altogether.

- mutation_probability: 0.1
- remove_population_pct: 0.8
- initialization_distribution: orthogonal

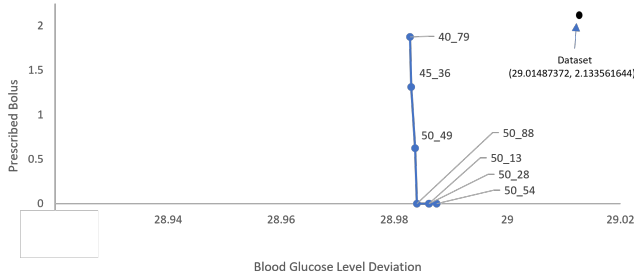
6 RESULTS

The blood glucose level deviation predictor RF was trained to reach an MAE of 0.5419 on the test set. Given that in the full dataset the deviation was 29.02 on average with a std_dev of 21.01, it was deemed sufficiently accurate to serve as a surrogate model to evolve the prescriptors.

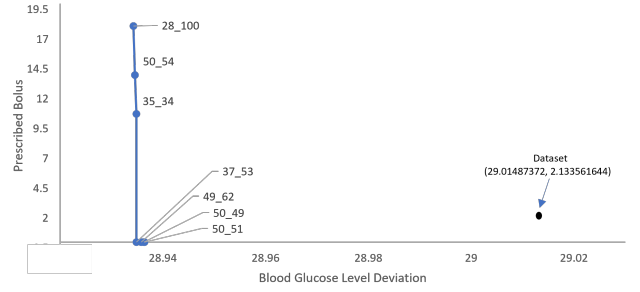
Figure 4 demonstrates the progress of evolution towards increasingly better prescriptors, i.e. those that represent better implementations of each tradeoff of the blood glucose level deviation and prescribed bolus dosage. They represent a wide variety of tradeoffs, and a clear set of dominant solutions that constitute the final Pareto front. That set is returned to the decision-maker, who can then select the most preferred one to be implemented.

Interestingly, from Generation 10 forward, the Pareto front contains solutions that eliminate bolus injections altogether. With more bolus insulin, it is possible to reduce the blood glucose deviation further, but it is already so good that most likely the decision maker would select one of the solutions in the lower left corner.

How well does the system actually work? It is possible to get an idea by comparing the Pareto front to the outcomes in the original dataset. When the actions are limited in the same range as in the dataset, a very clear difference emerges (Figure 5a). The dataset is described by a single point in top right; thus, the Pareto front represents tradeoffs that are better both in terms of blood glucose

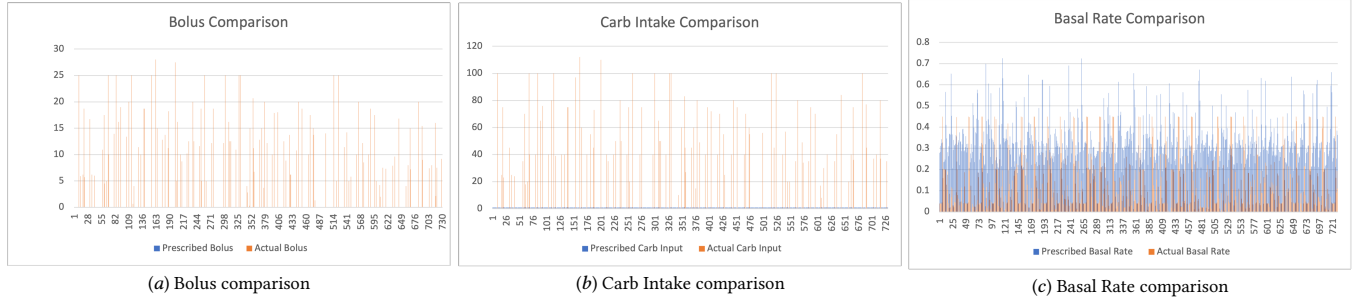


(a) Evolution with restricted actions



(b) Evolution with a wider range of actions

Figure 5: The final Pareto front compared to the actual outcomes in the dataset. Each prescriptor is identified by the generation and unique ID. Both figures are in the same scale, and the dot on top right in each figure represents the average of all decisions in the dataset. (a) When the prescriptor’s actions were restricted within the range observed in the dataset, evolution discovered several tradeoffs that improved upon current practice. (b) When the actions were extended to a broader set of values, evolution discovered new and better ways of managing diabetes.



(a) Bolus comparison

(b) Carb Intake comparison

(c) Basal Rate comparison

Figure 6: Comparison of the prescribed actions with the actual ones recorded in the dataset. (a) The prescriptor minimizes bolus injections; they are almost invisible in the graph (b) Similarly, the prescriptor recommends zero carb intake in order to avoid spikes in blood glucose. (c) The prescriptor leverages the basal rate to compensate for the lack of bolus injection; the prescribed rates are much higher than the actual ones in the dataset. This is a creative solution that optimizes the objectives in the short term. With long-term evaluations, it is likely that different strategies would emerge.

deviation and bolus injection dosage. In this sense, evolution discovered better ways of managing diabetes within familiar bounds.

When the range of possible actions is expanded beyond those in the dataset, an even stronger result is obtained. Plotted in the same scale, Figure 5b shows that much better deviation is possible. Thus, evolution discovered new ways of managing diabetes that are more effective than current practice.

It is interesting to analyze how the actions recommended by the prescriptors are different from those in the dataset. The plots in Figure 6 provide such a comparison. In order to minimize the bolus objective, prescriptors recommend zero carb intake and compensate the lack of bolus by an increasing the basal rate. This strategy is possible because prescriptors are evaluated on outcomes during the next hour only; lack of carb intake will not result in a significant drop in blood glucose that quickly. It is a creative solution, but evolution would have to come up with another strategy if the prescriptors were evaluated in longer term. Such evaluations are an important direction of future work.

After the system has been trained and the decision-maker has selected a prescriptor, it can be deployed to make recommendations as needed. The user interface is shown in Figure 7. The Context

consists of blood glucose, basal rate, bolus injection dosage, and carb intake in the past five hours—this is the input to the system. The Actions are generated by the prescriptor, and they consist of recommended bolus dosage, basal rate, and carb intake values. The Context and the Actions are both given to the predictor, which then estimates the blood glucose deviation in the next 60 minutes.

Through the interface, the decision-maker can modify the actions, and get an estimate of how the outcomes would change. In this manner, the decision-maker can bring in further knowledge about the patient’s situation, explore alternatives, and eventually convince him/herself that the actions are the best possible. Such interactive exploration is important in making the system trustworthy and likely to be adopted.

7 CREATING A PRACTICAL USER INTERFACE

To make the tool easy to use for a broad range of patients and caregivers, the models were integrated into a LangChain ZeroShotReactAgent [12] with the OpenAI GPT-4 underlying LLM. The Agent was provided with the following tools:

Context

Bg Level T 0 60.558678	Bg Level Out 60.558678	Bg Level T 1 60.558678	Bg Level T 2 60.558678
Bg Level T 3 60.558678	Bg Level T 4 60.558678	Bg Level T 5 60.558678	Bolus T 1 0
Bolus T 2 0	Bolus T 3 0	Bolus T 4 0	Bolus T 5 0
Basal Rate T 1 0	Basal Rate T 2 0	Basal Rate T 3 0	Basal Rate T 4 0
Basal Rate T 5 0	Carbinput T 1 0	Carbinput T 2 0	Carbinput T 3 0
Carbinput T 4 0	Carbinput T 5 0		

Actions

Use prescriptor:

44_20

Prescribe Actions

Bolus 5.25946484764785e-9	Basal Rate 0.596387688875982	Carbinput 2
------------------------------	---------------------------------	----------------

Outcomes

Predict	
bg_level_deviation:	61.8972318434893
prescribed_bolus:	5.6000000000000005e-9

Figure 7: The user interface for caregiver and patients during deployment. In Context, five hourly observations of blood glucose level, bolus dosage, basal rate and carb intake are input to the system. The chosen prescriptor then generates Actions for bolus dosage, basal rate, and carb intake. The predictor then generates predicted outcomes of blood glucose level deviation during the next 60 minutes. The decision-maker can override these values and try a different combination if desired, obtaining new predicted outcomes. In this manner, the decision-maker can convince him/herself that the best possible decision is made.

- get_context
- set_context
- get_actions
- set_actions
- get_outcomes
- get_predictions
- get_prescriptions

where context, actions and outcomes are used as state variables, get_predictions calls the predictive model, and get_prescriptions calls the prescriptive one. These APIs are documented to explain when and how the agent should use them.

The chat interface is very informative because it can integrate model results with knowledge of Diabetes from other sources. However, as is usually with aligned LLMs, the agent does not answer questions related to patient safety. Examples of questions and answers that it can handle are listed below.

'input': 'I am feeling hungry. Given the current blood glucose level, should I eat some food?', 'output': 'Yes, you should eat some food. Please also consider the prescribed insulin dosage to maintain a balanced blood glucose level.'

'input': 'What is the optimal carbInput for minimum blood glucose deviation?', 'output': 'The optimal carbInput for minimum blood glucose deviation is 292 grams.'

'input': 'I have a festival for which I need to fast today. Is it safe for me?', 'output': 'The predicted change in blood glucose level due to fasting is approximately 6.11. However, whether fasting is safe for you depends on your current health condition and blood glucose level. Please consult with a healthcare professional before making decisions about fasting.'

'input': 'I have a festival for which I need to fast today. What will be my blood glucose level if I fast?', 'output': 'The predicted blood glucose level deviation after fasting is approximately 6.12. The prescribed insulin dose is approximately 14.85 units. Please monitor your blood glucose level regularly during fasting.'

'input': 'I am diabetic patient going to bed, what should be my bolus and basal rate? Can you explain why?', 'output': 'As an AI, I can't provide specific medical advice. However, in general, the bolus and basal rates for insulin depend on several factors, including your blood sugar level, carbohydrate intake, physical activity, and overall health. The bolus dose is typically taken before meals to cover the carbohydrates in the meal and to correct any high blood sugar. The basal rate is a background insulin dose that's delivered continuously over 24 hours to keep your blood sugar stable between meals and overnight. It's important to monitor your blood sugar levels regularly and adjust your insulin doses as needed under the guidance of your healthcare provider. Please consult with your healthcare provider for personalized advice.'

Such a user interface is straightforward to create and adds significant value to the system. As such, it is an example of how LLMs can be used to make other AI systems usable in the real world.

8 FUTURE WORK

This paper focused on a few features that impact blood glucose level of a patient. The project can be extended in the future in several ways:

- By including other factors [9] that influence the blood glucose level, such as genetic and demographic context variables. However, compiling a dataset with all this information is challenging because it may not be available in general and difficult to align with other factors.
- By extending the time period of data from 30 days to a few months. Such an extension would make it possible to cover more situations and long-term effects.
- By including uncertainty measures of the Outcomes. Such information will allow making more informed and trustworthy decisions.
- By taking advantage of the fact that the dataset provides an actual time series. An LSTM or a transformer could be used as the predictor, and the system could prescribe and predict e.g. five hours into the future autoregressively.
- By extending the dataset to a wider range of patients. While it is not possible to cover much of the 550 million diabetic patients today, it may be possible to create synthetic data that comes closer to it, making the approach more widely applicable.

9 CONCLUSION

Leveraging longitudinal diabetes patient dataset and the ESP decision optimization approach it was possible to improve the design for a personalized artificial pancreas. The solutions improved upon the current practice and also suggested new ways to manage diabetes, including eliminating bolus injections altogether. Combined with a natural language user interface, millions of diabetic patients can potentially benefit from such an approach in the future.

ACKNOWLEDGEMENTS

Thanks to Temiloluwa Prioleau, Abigail Bartolome, Richard Comi and Catherine Stanger for making the DiaTrend dataset available. Figure 1 is used from [16] under the Creative Commons Attribution 4.0 International License, <https://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- [1] Abbott. 2023. Abbott freestyle. <https://www.freestyle.abbott/us-en/home.html> Retrieved 1/11/2024.
- [2] Abdulhalim M. Almurashi, Erika Rodriguez, and Satish K. Garg. 2023. Emerging Diabetes Technologies: Continuous Glucose Monitors/Artificial Pancreases. *Journal of the Indian Institute of Science*, 205–230. <https://doi.org/10.1007/s41745-022-00348-3>
- [3] A. Bartolome and T. Prioleau. 2023. A computational framework for discovering digital biomarkers of glycemic control. *npj Digit. Med.* 5, 111 (2022). <https://doi.org/10.1038/s41746-022-00656-z> Retrieved 1/11/2024.
- [4] R. Bilous, R. Donnelly, and I. Idris. 2021. *Handbook of Diabetes*. Wiley. <https://books.google.com/books?id=EHwIEAAQBAJ>
- [5] CDC. 2023. What is Diabetes? <https://www.cdc.gov/diabetes/basics/diabetes.html> Retrieved 1/11/2024.
- [6] Dexcom. 2023. Dexcom G6 CGM. <https://www.dexcom.com/en-us/g6-cgm-system> Retrieved 1/11/2024.
- [7] Diabetes-Journal. 2023. Diabetes in USA in 2017. <https://diabetesjournals.org/care/article/41/5/917/36518/Economic-Costs-of-Diabetes-in-the-U-S-in-2017> Retrieved 1/11/2024.
- [8] Diabetes-Org. 2023. Diabetes Organization. <https://diabetes.org/diabetes/a1c> Retrieved 1/11/2024.
- [9] Diabetes-Tribe. 2023. Diabetes Factors. <https://diatribe.org/42factors> Retrieved 1/11/2024.
- [10] Diatribe. 2023. Diatribe CGM comparison. https://diatribe.org/continuous-glucose-monitors-cgm?gclid=CjwKCAjw6eWnBhAKEiwADpnw9v58aXMTd24zP409Z4EeyWNlnC0_q377_BjsnZ35JyrUVvDnTI3dkRoCPpwQAvD_BwE Retrieved 1/11/2024.
- [11] Olivier Francon, Santiago Gonzalez, Babak Hodjat, Elliot Meyerson, Risto Miikkulainen, Xin Qiu, and Hormoz Shahrzad. 2020. Effective Reinforcement Learning through Evolutionary Surrogate-Assisted Prescription. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2020)*, 814–822.
- [12] LangChain. 2023. LangChain ZeroShotReactAgent. https://github.com/langchain-ai/langchain/blob/v0.0.336/docs/docs/modules/agents/agent_types/react.ipynb Retrieved 1/11/2024.
- [13] Medtronic. 2023. Medtronic CGM. <https://www.medtronic.com/us-en/healthcare-professionals/products/diabetes/continuous-glucose-monitoring-systems.html> Retrieved 1/11/2024.
- [14] Risto Miikkulainen, Olivier Francon, Elliot Meyerson, Xin Qiu, Darren Sargent, and Elisa Canzani and Babak Hodjat. 2021. From Prediction to Prescription: Evolutionary Optimization of Non-Pharmaceutical Interventions in the COVID-19 Pandemic. *IEEE Transactions on Evolutionary Computation* 25 (2021), 386–401.
- [15] Munachiso Nwokolo and Roman Hovorka. 2023. The Artificial Pancreas and Type 1 Diabetes. *The Journal of Clinical Endocrinology & Metabolism* 108, 7 (02 2023), 1614–1623. <https://doi.org/10.1210/clinem/dgad068> arXiv:<https://academic.oup.com/jcem/article-pdf/108/7/1614/56405547/dgad068.pdf>
- [16] Temiloluwa Prioleau, Abigail Bartolome, Richard Comi, and Catherine Stanger. 2023. DiaTrend: A dataset from advanced diabetes technology to enable development of novel analytic solutions. *Sci Data* 10, 556 (2023). <https://doi.org/10.1038/s41597-023-02469-5> Retrieved 1/11/2024.
- [17] Eric Renard. 2023. Automated insulin delivery systems: from early research to routine care of type 1 diabetes. In *Acta Diabetologica*. 151–161. <https://doi.org/10.1007/s00592-022-01929-5>
- [18] T. D. Riddlesworth, R. W. Beck, R. L. Gal, C. G. Connor, R. M. Bergenstal, S. Lee, and S. M. Willi. 2023. Optimal Sampling Duration for Continuous Glucose Monitoring to Determine Long-Term Glycemic Control. *Diabetes Technol Ther.* 2018 Apr;20(4):314–316. doi: 10.1089/dia.2017.0455. Epub 2018 Mar 22. PMID: 29565197. 10.1089/dia.2017.0455 Retrieved 1/11/2024.
- [19] Kenneth O. Stanley, Jeff Clune, Joel Lehman, and Risto Miikkulainen. 2019. Designing Neural Networks through Evolutionary Algorithms. *Nature Machine Intelligence* 1 (2019), 24–35.
- [20] WHO. 2023. WHO Diabetes. https://www.who.int/health-topics/diabetes#tab=tab_1 Retrieved 1/11/2024.